Information theoretical properties of Tsallis entropies

Shigeru FURUICHI[1][*]

[1]Department of Electronics and Computer Science,
Tokyo University of Science, Yamaguchi, 756-0884, Japan

Abstract. A chain rule and a subadditivity for the entropy of type β, which is one of the nonadditive entropies, were derived by Z.Daróczy. In this paper, we study the further relations among Tsallis type entropies which are typical nonadditive entropies. The chain rule is generalized by showing it for Tsallis relative entropy and the nonadditive entropy. We show some inequalities related to Tsallis entropies, especially the strong subadditivity for Tsallis type entropies and the subadditivity for the nonadditive entropies. The subadditivity and the strong subadditivity naturally lead to define Tsallis mutual entropy and Tsallis conditional mutual entropy, respectively, and then we show again chain rules for Tsallis mutual entropies. We give properties of entropic distances in terms of Tsallis entropies. Finally we show parametrically extended results based on information theory.

1 Introduction

For the probability distribution $p(x) \equiv p(X = x)$ of the random variable X, Tsallis entropy was defined in [1] :

$$S_q(X) \equiv -\sum_x p(x)^q \ln_q p(x), \quad (q \neq 1) \tag{1}$$

with one parameter q as an extension of Shannon entropy, where q-logarithm function is defined by $\ln_q(x) \equiv \frac{x^{1-q}-1}{1-q}$ for any nonnegative real number x and q. Tsallis entropy plays an essential role in nonextensive statistics so that many important results have been published in mainly statistical physics [2]. In particular, see [3, 4] for the recent results on Tsallis entropies. The field of the study on Tsallis entropy and related fields in statistical physics are often called Tsallis statistics or nonextensive statistical physics due to the nonextensivity Eq.(3) in the below. We easily find that $\lim_{q \to 1} S_q(X) = S_1(X) \equiv -\sum_x p(x) \log p(x)$, since q-logarithm converges to natural logarithm as $q \to 1$. After the celebrated discover by Tsallis, it was shown that Tsallis entropy can be uniquely formulated by the generalized Shannon-Khinchin's axioms in [5, 6, 7]. The axiomatical characterization of Tsallis entropy was improved and that of Tsallis relative entropy was formulated in [8]. It is a crucial difference of the fundamental property between Shannon entropy and Tsallis entropy that for two independent random variables X and Y,

---

[*]E-mail:furuichi@ed.yama.tus.ac.jp

Shannon entropy $S_1(X)$ has an additivity :

$$S_1(X \times Y) = S_1(X) + S_1(Y), \qquad (2)$$

however, Tsallis entropy $S_q(X), (q \neq 1)$ has a pseudoadditivity :

$$S_q(X \times Y) = S_q(X) + S_q(Y) + (1 - q)S_q(X)S_q(Y), \qquad (3)$$

where $S_q(X \times Y)$ is Tsallis joint entropy, which will be defined in the below, for the direct product $X \times Y$ of two independent random variables X and Y. The independence between X and Y phyisically means that there is no interaction between two systems X and Y. This pseudoadditivity of Tsallis entropy in Eq.(3) originally comes from that of q-logarithm function $\ln_q(x)$. We pay attention whether an additivity (such as Eq.(2)) of entropies for two independent random variables holds or not (such as Eq.(3)). An entropy is called a nonadditive entropy if it does not satisfy the additivity such as Eq.(2), while an entropy is called an additive entropy if it satisfies the additivity specified in Eq.(2). The entropy of type β introduced by Daróczy in [17] (also called the structural a-entropy introduced by Havrda and Charvát in [18]):

$$H_\beta(X) \equiv \frac{\sum_x \left(p(x)^\beta - p(x)\right)}{2^{1-\beta} - 1}, \quad (\beta > 0, \beta \neq 1)$$

is a nonadditive entropy.

We are now interested in Tsallis entropy which is a typical nonadditive entropy. From Eq.(3), we easily find that we have a subadditivity of Tsallis entropies for two independent random variables X and Y in the case of $q \geq 1$ since Tsallis entropy is nonnegative. In the first of the section 3, we will review a subadditivity of Tsallis entropies for general random variables X and Y in the case of $q \geq 1$.

From the entropy theoretical point of view, for the related entropies such as Tsallis conditional entropy and Tsallis joint entropy were introduced in [9, 10, 11] and then some interesting relations between them were derived. In this paper, in order to derive the relation between them, we adopt the following definition of Tsallis conditional entropy and Tsallis joint entropy.

Definition 1.1 For the conditional probability $p(x|y) \equiv p(X = x|Y = y)$ and the joint probability $p(x, y) \equiv p(X = x, Y = y)$, we define Tsallis conditional entropy and Tsallis joint entropy by

$$S_q(X|Y) \equiv - \sum_{x,y} p(x, y)^q \ln_q p(x|y), \quad (q \neq 1), \qquad (4)$$

and

$$S_q(X, Y) \equiv - \sum_{x,y} p(x, y)^q \ln_q p(x, y), \quad (q \neq 1). \qquad (5)$$

We note that the above definitions were essentially introduced in [17] by

$$H_\beta(X, Y) \equiv \frac{\sum_{x,y} \left(p(x, y)^\beta - p(x, y)\right)}{2^{1-\beta} - 1}, \quad (\beta > 0, \beta \neq 1)$$

$$H_\beta(X|Y) \equiv \sum_y p(y)^\beta H_\beta(X|y), \quad (\beta > 0, \beta \neq 1)$$

except for the difference of the multiplicative function. And then a chain rule and a subadditivity:

$$H_\beta(X, Y) = H_\beta(X) + H_\beta(Y|X),$$
$$H_\beta(Y|X) \leq H_\beta(Y), \quad \beta > 1,$$

were shown in Theorem 8 of [17]. In this paper, we call the entropy defined by Eq.(1), Eq.(4) or Eq.(5) Tsallis type entropy. It is remarkable that Eq.(4) can be easily deformed by

$$S_q(X|Y) = \sum_y p(y)^q S_q(X|y). \tag{6}$$

Note that for $q \neq 1$ we have $S_q(X|Y) = \sum_y p(y)^q S_q(X) \neq S_q(X)$ when X and Y are indepentent each other (i.e., $p(x|y) = p(x)$ for all x and y). Eq.(6) can be further regarded that the following expectation value :

$$E_q(X) \equiv \sum_i p_i^q x_i$$

depending on the parameter q is adopted. Since the expectation value $E_q(X)$ has a crucial lack that it dose not hold $E_q(1) = 1$ for $q \neq 1$ in general, a slightly modified definitions, the normalized Tsallis entropies such as

$$S_q^{nor}(X) \equiv \frac{-\sum_x p(x)^q \ln_q p(x)}{\sum_x p(x)^q}, \quad (q \neq 1) \tag{7}$$

are sometimes used [10, 11], and then important and meaningful results in physics are derived for the normalized nonadditive conditional entropy defined in [6, 10]. However, as pointed out in [13], the normalized Tsallis entropies are not stable, but the original Tsallis entropies are stable as experimental quantities so that it is meaningful to study the original Tsallis entropies from the physical point of view. Inspired by these pioneering works [6, 10, 13], we therefore study the information theoretical properties of the original Tsallis type entropies without normalization in this paper. As advantages in information theoretical study, our approach shows chain rules and subadditivities of the Tsallis entropies and then naturally leads to define the Tsallis mutual entropy. Finally the rough meaning of the parameter q of the Tsallis entropy is given from the information theoretical viewpoint. Both entropies defined in Eq.(4) and Eq.(5) recover the usual Shannon entropies as $q \to 1$.

In the past, some parametrically extended entropies were introduced in the mathematical point of view. Especially, the Rényi entropy [14, 15] (see also [16])

$$R_q(X) = \frac{1}{1-q} \log \sum_x p(x)^q, \quad (q \neq 1). \tag{8}$$

is famous. For Tsallis entropy and Rényi entropy, we have the following relation,

$$R_q(X) = \frac{1}{1-q} \log \{1 + (1-q) S_q(X)\}, \quad (q \neq 1). \tag{9}$$

It is notable that Rényi entropy has an additivity:

$$R_q(X \times Y) = R_q(X) + R_q(Y), \tag{10}$$

for two independent random variables X and Y. Although there exsits a simple transformation between Tsallis entropy and Rényi entropy, they have a different structure in regards to the matter of their additivities.

This paper is organized as follows. In section 2, we review a chain rule given in [17] and give its generalized results. We also give the chain rule for Tsallis relative entropy. In section 3, we review a subadditivity given in [17] and give its generalized results. In addition, we show several inequalities related Tsallis entropies. In section 4, we introduce Tsallis mutual entropy and then give its fundamental properties. In section 5, we discuss on entropic distances and correlation coefficients due to Tsallis entropies. In section 6, we give some applications of Tsallis entropies in the context of information theory. Related remarks and discussions are given in section 7.

2   Chain rules for Tsallis entropies

It is important to study so-called a chain rule which gives the relation between a conditional entropy and a joint entropy in not only information theory [22] but also statistical physics. For these Tsallis entropies, the following chain rule holds as similar as the chain rule holds for the joint entropy of type β and the conditional entropy of type β.

Proposition 2.1 ([17])

$$S_q(X,Y) = S_q(X) + S_q(Y|X). \tag{11}$$

(Therefore immediately $S_q(X) \le S_q(X,Y)$.)

(Proof) From the formulas $\ln_q(xy) = \ln_q x + x^{1-q} \ln_q y$ and $\ln_q \frac{1}{x} = -x^{q-1} \ln_q x$, we have

$$
\begin{aligned}
S_q(X,Y) &= \sum_x \sum_y p(x,y) \ln_q \frac{1}{p(x,y)} = \sum_x \sum_y p(x,y) \ln_q \frac{1}{p(x)p(y|x)} \\
&= \sum_x p(x) \ln_q \frac{1}{p(x)} + \sum_x \sum_y p(x,y) p(x)^{q-1} \ln_q \frac{1}{p(y|x)} \\
&= -\sum_x p(x)^q \ln_q p(x) - \sum_x p(x)^q \sum_y p(y|x)^q \ln_q p(y|x) \\
&= S_q(X) + S_q(Y|X).
\end{aligned}
$$

∎

From the process of the proof of Proposition 2.1, we can easily find that if X and Y are independent (i.e., $p(y|x) = p(y)$ for all x and y), then Eq.(11) recovers the pseudoadditivity Eq.(3):

$$
\begin{aligned}
S_q(X,Y) &= -\sum_x p(x)^q \ln_q p(x) - \sum_x p(x)^q \sum_y p(y)^q \ln_q p(y) \\
&= S_q(X) + \sum_x p(x)^q S_q(Y) \\
&= S_q(X) + S_q(Y) + (1-q) S_q(X) S_q(Y).
\end{aligned}
$$

As a corollary of the above Proposition 2.1, we have the following lemma.

Lemma 2.2 The following chain rules hold.

(1) $S_q(X,Y,Z) = S_q(X,Y|Z) + S_q(Z)$.

(2) $S_q(X,Y|Z) = S_q(X|Z) + S_q(Y|X,Z)$.

(Proof) By using $\ln_q \frac{y}{x} = \ln_q y + y^{1-q} \ln_q \frac{1}{x}$, we have

$$S_q(X,Y|Z) = S_q(X,Y,Z) - S_q(Z).$$

In the similar way, we have

$$S_q(Y|X,Z) = S_q(X,Y,Z) - S_q(X,Z).$$

Also from the Proposition 2.1, we have

$$S_q(X,Y) = S_q(X) + S_q(Y|X).$$

Therefore we have

$$
\begin{aligned}
S_q\left(Y \mid X, Z\right) &= S_q\left(X, Y, Z\right) - S_q\left(X, Z\right) \\
&= S_q\left(X, Y \mid Z\right) + S_q\left(Z\right) - \left\{S_q\left(X \mid Z\right) + S_q\left(Z\right)\right\} \\
&= S_q\left(X, Y \mid Z\right) - S_q\left(X \mid Z\right).
\end{aligned}
$$

∎

From (2) of of Lemma 2.2, we have

$$
S_q(X|Z) \leq S_q(X, Y|Z) \tag{12}
$$

Remark 2.3 (2) of Lemma 2.2 can be generalized in the following formula :

$$
S_q\left(X_1, \cdots, X_n \mid Y\right) = \sum_{i=1}^{n} S_q\left(X_i \mid X_{i-1}, \cdots, X_1, Y\right). \tag{13}
$$

Theorem 2.4 Let $X_1, X_2, \cdots, X_n$ be the random variables obeying to the probability distribution $p(x_1, x_2, \cdots, x_n)$. Then we have the following chain rule.

$$
S_q\left(X_1, X_2, \cdots, X_n\right) = \sum_{i=1}^{n} S_q\left(X_i \mid X_{i-1}, \cdots, X_1\right). \tag{14}
$$

(Proof) We prove the theorem by induction on n. From Proposition 2.1, we have

$$
S_q\left(X_1, X_2\right) = S_q\left(X_1\right) + S_q\left(X_2 \mid X_1\right).
$$

Assuming that Eq.(14) holds for some n, from Eq.(11) we have

$$
\begin{aligned}
S_q\left(X_1, \cdots, X_{n+1}\right) &= S_q\left(X_1, \cdots, X_n\right) + S_q\left(X_{n+1} \mid X_1, \cdots, X_n\right) \\
&= \sum_{i=1}^{n} S_q\left(X_i \mid X_{i-1}, \cdots, X_1\right) + S_q\left(X_{n+1} \mid X_n, \cdots, X_1\right),
\end{aligned}
$$

which means Eq.(14) holds for $n + 1$. ∎

In the rest of this section, we discuss on the chain rule for Tsallis relative entropy. To do so, we use Tsallis relative entropy

$$
D_q(U|V) \equiv -\sum_{x} u(x) \ln_q \frac{v(x)}{u(x)}, \tag{15}
$$

defined for two probability distributions $u(x)$ and $v(x)$, and $q \geq 0$. See [19, 20, 21, 23] for details of Tsallis relative entropy. Eq.(15) is equivalent to the following forms,

$$
\begin{aligned}
D_q(U|V) &= \frac{\sum_{x} \left(u(x) - u(x)^q v(x)^{1-q}\right)}{1-q} \tag{16} \\
&= \sum_{x} u(x)^q v(x)^{1-q} \ln_q \frac{u(x)}{v(x)}. \tag{17}
\end{aligned}
$$

Along the line of the definition of Tsallis conditional entropy Eq.(4), we naturally define Tsallis conditional relative entropy such as

$$
\begin{aligned}
D_q\left(u\left(y|x\right)|v\left(y|x\right)\right) &\equiv \sum_{x,y} u(x,y)^q v(x,y)^{1-q} \ln_q \frac{u(y|x)}{v(y|x)} \qquad (18)\\
&= \sum_x u(x)^q v(x)^{1-q} \sum_y u(y|x)^q v(y|x)^{1-q} \ln_q \frac{u(y|x)}{v(y|x)} \qquad (19)
\end{aligned}
$$

for two joint probability distributions $u(x,y)$ and $v(x,y)$, and two conditional probability distributions $u(y|x)$ and $v(y|x)$, based on the view of Eq.(17), neither Eq.(15) nor Eq.(16). Taking the limit as $q \to 1$, Eq.(18) recovers the usual conditional relative entropy. Then we have the following chain rule of Tsallis relative entropy for general (possibly *not* independent) case.

Theorem 2.5 (Chain rule for Tsallis relative entropy)

$$
D_q\left(u\left(x,y\right)|v\left(x,y\right)\right) = D_q\left(u\left(x\right)|v\left(x\right)\right) + D_q\left(u\left(y|x\right)|v\left(y|x\right)\right). \qquad (20)
$$

(Proof) The proof follows by the direct calculations.

$$
\begin{aligned}
D_q\left(u\left(x,y\right)|v\left(x,y\right)\right) &= -\sum_{x,y} u\left(x,y\right) \ln_q \frac{v\left(x,y\right)}{u\left(x,y\right)} = -\sum_{x,y} u\left(x,y\right) \ln_q \frac{v\left(x\right) v\left(y|x\right)}{u\left(x\right) u\left(y|x\right)}\\
&= -\sum_{x,y} u\left(x,y\right) \left\{ \ln_q \frac{v\left(x\right)}{u\left(x\right)} + \ln_q \frac{v\left(y|x\right)}{u\left(y|x\right)} + (1-q) \ln_q \frac{v\left(x\right)}{u\left(x\right)} \ln_q \frac{v\left(y|x\right)}{u\left(y|x\right)} \right\}\\
&= -\sum_{x,y} u\left(x,y\right) \left\{ \ln_q \frac{v\left(x\right)}{u\left(x\right)} + \left\{ \frac{u\left(x\right)}{v\left(x\right)} \right\}^{q-1} \ln_q \frac{v\left(y|x\right)}{u\left(y|x\right)} \right\}\\
&= D_q\left(u\left(x\right)|v\left(x\right)\right) - \sum_{x,y} \left\{ \frac{u\left(x\right)}{v\left(x\right)} \right\}^{q-1} u\left(x,y\right) \ln_q \frac{v\left(y|x\right)}{u\left(y|x\right)}.\\
&= D_q\left(u\left(x\right)|v\left(x\right)\right) + D_q\left(u\left(y|x\right)|v\left(y|x\right)\right).
\end{aligned}
$$

∎

Taking the limit as $q \to 1$, Theorem 2.5 recovers Theorem 2.5.3 in [22]. It is known that for $u\left(y|x\right) = u\left(y\right)$ and $v\left(y|x\right) = v\left(y\right)$, Tsallis relative entropy has a pseudoadditivity :

$$
D_q\left(u(x)u(y)|v(x)v(y)\right) = D_q\left(u(x)|v(x)\right) + D_q\left(u(y)|v(y)\right) + (q-1) D_q\left(u(x)|v(x)\right) D_q\left(u(y)|v(y)\right). \qquad (21)
$$

From the process of the above proof, we find that the chain rule Eq.(20) recovers the pseudoadditivity Eq.(21) when $u\left(y|x\right) = u\left(y\right)$ and $v\left(y|x\right) = v\left(y\right)$.

3 Subadditivities for Tsallis entropies

From Eq.(3), for $q \geq 1$ and two independent random variables X and Y, the subadditivity holds:

$$
S_q(X \times Y) \leq S_q(X) + S_q(Y).
$$

It is known that the subadditivity for general random variables X and Y holds in the case of $q \geq 1$, thanks to the following proposition.

Proposition 3.1 ([17]) The following inequality holds for two random variables X and Y, and $q \geq 1$,

$$S_q(X|Y) \leq S_q(X), \tag{22}$$

with equality if and only if $q = 1$ and $p(x|y) = p(x)$ for all x and y.

(Proof) We give the similar proof given in [17] as a version of Tsallis type entropies for the convenience of the reader. We define the function $\mathrm{Ln}_q(x) \equiv \frac{x^q - x}{1-q}$ for $q \geq 0, q \neq 1$ and $x \geq 0$. Then the function Ln_q is nonnegative for $0 \leq x \leq 1, q > 1$ and concave in x for $q > 0, q \neq 1$. By the concavity of Ln_q, we have

$$\sum_y p(y)\mathrm{Ln}_q(p(x|y)) \leq \mathrm{Ln}_q\left(\sum_y p(y)p(x|y)\right).$$

Taking the summation of the both sides on the above inequality on x, we have

$$\sum_y p(y)\sum_x \mathrm{Ln}_q(p(x|y)) \leq \sum_x \mathrm{Ln}_q(p(x)) \tag{23}$$

Since $p(y)^q \leq p(y)$ for any y and $q > 1$, and $\mathrm{Ln}_q(t) \geq 0$ for any $t \geq 0$ and $q > 1$, we have,

$$p(y)^q \sum_x \mathrm{Ln}_q(p(x|y)) \leq p(y)\sum_x \mathrm{Ln}_q(p(x|y)).$$

Taking the summation of the both sides on the above inequality on y, we have

$$\sum_y p(y)^q \sum_x \mathrm{Ln}_q(p(x|y)) \leq \sum_y p(y)\sum_x \mathrm{Ln}_q(p(x|y)). \tag{24}$$

By Eq.(23) and Eq.(24), we have

$$\sum_y p(y)^q \sum_x \mathrm{Ln}_q(p(x|y)) \leq \sum_x \mathrm{Ln}_q(p(x)).$$

Therefore we have $S_q(X|Y) \leq S_q(X)$. The equality holds when $q = 1$ and $p(x|y) = p(x)$, which means X and Y are independent each other. Thus the proof of the proposition was completed. ∎

Eq.(22) and Eq.(11) imply the subadditivity of Tsallis entropies.

Theorem 3.2 ([17]) For $q \geq 1$, we have

$$S_q(X,Y) \leq S_q(X) + S_q(Y). \tag{25}$$

From Eq.(22) and Theorem 2.4, we have the following extended relation which is often called *independence bound on entropy* in information theory.

Theorem 3.3 Let $X_1, X_2, \cdots, X_n$ be the random variables obeying to the probability distribution $p(x_1, x_2, \cdots, x_n)$. Then for $q \geq 1$, we have

$$S_q(X_1, \cdots, X_n) \leq \sum_{i=1}^n S_q(X_i) \tag{26}$$

with equality if and only if $q = 1$ the random variables are independent each other.

On the other hand, we easily find that for two independent random variables X and Y, and $0 \leq q < 1$, the superadditivity holds:

$$S_q(X \times Y) \geq S_q(X) + S_q(Y).$$

However, in general the superadditivity for two correlated random variables X and Y, and $0 \leq q < 1$ does not hold. Becasue we can show many counterexamples. For example, we consider the following joint distribution of X and Y.

$$p(x_1, y_1) = p(1 - x), p(x_1, y_2) = (1 - p)y, p(x_2, y_1) = px, p(x_2, y_2) = (1 - p)(1 - y), \quad (27)$$

where $0 \leq p, x, y \leq 1$. Then each marginal distribution can be computed by

$$p(x_1) = p(1 - x) + (1 - p)y, p(x_2) = px + (1 - p)(1 - y), p(y_1) = p, p(y_2) = 1 - p. \quad (28)$$

In general, we clearly see X and Y are not independent each other for the above example. Then the value of $\Delta \equiv S_q(X, Y) - S_q(X) - S_q(Y)$ takes both positive and negative so that there does not exist the complete ordering between $S_q(X, Y)$ and $S_q(X) + S_q(Y)$ for correlated X and Y in the case of $0 \leq q < 1$. Indeed, $\Delta = -0.287089$ when $q = 0.8, p = 0.6, x = 0.1, y = 0.1$, while $\Delta = 0.0562961$ when $q = 0.8, p = 0.6, x = 0.1, y = 0.9$.

Up to the above discussions, we have the possibility that the strong subadditivity holds in the case of $q \geq 1$. Indeed we can prove the following strong subadditivity for Tsallis type entropies.

Theorem 3.4 For $q \geq 1$, the strong subadditivity

$$S_q(X, Y, Z) + S_q(Z) \leq S_q(X, Z) + S_q(Y, Z) \quad (29)$$

holds with equality if and only if $q = 1$ and, random variables X and Y are independent for a given random variable Z.

(Proof) Theorem is proven by the similar way of Proposition 3.1. By the concavity of Ln_q, we have

$$\sum_y p(y|z)\mathrm{Ln}_q(p(x|y, z)) \leq \mathrm{Ln}_q\left(\sum_y p(y|z)p(x|y, z)\right)$$

Multiplying $p(z)^q$ to the both sides of the above inequality and then taking the summation on z and x, we have

$$\sum_{y,z} p(z)^q p(y|z) \sum_x \mathrm{Ln}_q(p(x|y, z)) \leq \sum_z p(z)^q \sum_x \mathrm{Ln}_q(p(x|z)) \quad (30)$$

since $\sum_y p(y|z)p(x|y, z) = p(x|z)$.

For any y, z and $q \geq 1$, we have $p(y|z)^q \leq p(y|z)$. Then by the nonnegativity of the function Ln_q, we have

$$p(y|x)^q \sum_x \mathrm{Ln}_q(p(x|y, z)) \leq p(y|z) \sum_x \mathrm{Ln}_q(p(x|y, z)).$$

Multiplying $p(z)^q$ to the both sides of the above inequality and then taking the summation on y and z, we have

$$\sum_{y,z} p(z)^q p(y|x)^q \sum_x \mathrm{Ln}_q(p(x|y, z)) \leq \sum_{y,z} p(z)^q p(y|z) \sum_x \mathrm{Ln}_q(p(x|y, z)). \quad (31)$$

From Eq.(30) and Eq.(31), we have

$$\sum_{y,z} p(y,z)^q \sum_x \mathrm{Ln}_q(p(x|y,z)) \leq \sum_z p(z)^q \sum_x \mathrm{Ln}_q(p(x|z)).$$

Thus we have

$$S_q\left(X\,|Y,Z\right) \leq S_q\left(X\,|Z\right), \tag{32}$$

which implies

$$S_q\left(X,Y,Z\right) - S_q\left(Y,Z\right) \leq S_q\left(X,Z\right) - S_q\left(Z\right),$$

by chain rules. The equality holds when $q = 1$ and $p(x|y,z) = p(x|z)$, which means X and Y are independent each other for a given Z. ∎

We find that Eq.(29) recovers Eq.(25), by taking the random variable Z as a trivial one. This means that Eq.(29) is a generalization of Eq.(25). We have the further generalized inequality by the similar way of the proof of Theorem 3.4.

Theorem 3.5 Let $X_1, \cdots, X_{n+1}$ be the random variables. For $q > 1$, we have

$$S_q(X_{n+1}|X_1, \cdots, X_n) \leq S_q(X_{n+1}|X_2, \cdots, X_n). \tag{33}$$

The subadditivity for Tsallis entropies conditioned by Z holds.

Proposition 3.6 For $q \geq 1$, we have

$$S_q(X,Y|Z) \leq S_q(X|Z) + S_q(Y|Z). \tag{34}$$

(Proof) Summing $-2S_q(Z)$ to the both sides in Eq.(29), we have

$$S_q(X,Y,Z) - S_q(Z) \leq S_q(X,Z) - S_q(Z) + S_q(Y,Z) - S_q(Z).$$

By chain rules, we have the proposition. ∎

Proposition 3.6 can be generalized in the following.

Theorem 3.7 For $q \geq 1$, we have

$$S_q(X_1, \cdots, X_n|Z) \leq S_q(X_1|Z) + \cdots + S_q(X_n|Z). \tag{35}$$

In addition, we have the following propositions.

Proposition 3.8 For $q \geq 1$, we have

$$2S_q(X,Y,Z) \leq S_q(X,Y) + S_q(Y,Z) + S_q(Z,X).$$

(Proof) Form Eq.(22) and Eq.(32), we have

$$S_q(X|Y,Z) \leq S_q(X),$$

and

$$S_q(Z|X,Y) \leq S_q(Z|X).$$

Thus we have

$$S_q(X|Y,Z) + S_q(Z|X,Y) \leq S_q(X) + S_q(Z|X).$$

By chain rules, we have

$$S_q(X,Y,Z) - S_q(X,Y) + S_q(X,Y,Z) - S_q(Y,Z) \leq S_q(Z,X),$$

which implies the proposition. ∎

Proposition 3.9 For $q > 1$, we have

$$S_q(X_n|X_1) \leq S_q(X_2|X_1) + \cdots + S_q(X_n|X_{n-1}).$$

(Proof) From Eq.(14) and Eq.(32), we have

$$S_q(X_1, \cdots, X_n) \leq S_q(X_1) + S_q(X_2|X_1) + \cdots S_q(X_n|X_{n-1}).$$

Therefore we have

$$
\begin{aligned}
S_q(X_n|X_1) &\leq S_q(X_2, \cdots, X_n|X_1) \\
&= S_q(X_1, \cdots, X_n) - S_q(X_1) \\
&\leq S_q(X_2|X_1) + \cdots + S_q(X_n|X_{n-1}),
\end{aligned}
$$

by Eq.(11) and Eq.(12). ∎

4 Tsallis mutual entropy

For normalized Tsallis entropies, the mutual information was defined in [11] with the assumption of its nonegativity. We define the Tsallis mutual entropy in terms of the original (*not* normalized) Tsallis type entropies. The inequality Eq.(22) naturally leads us to define Tsallis mutual entropy without the assumption of its nonegativity.

Definition 4.1 For two random variables X and Y, and $q > 1$, we define the Tsallis mutual entropy as the difference between Tsallis entropy and Tsallis conditional entropy such that

$$I_q(X;Y) \equiv S_q(X) - S_q(X|Y). \tag{36}$$

Note that we never use the term *mutual information* but use mutual entropy through this paper, since a proper evidence of channel coding theorem for information transmission has not ever been shown in the context of Tsallis statistics. From Eq.(11), Eq.(25) and Eq.(22), we easily find that $I_q(X;Y)$ has the following fundamental properties.

Proposition 4.2   (1) $0 \leq I_q(X;Y) \leq min\{S_q(X), S_q(Y)\}$.

(2) $I_q(X;Y) = S_q(X) + S_q(Y) - S_q(X,Y) = I_q(Y;X)$.

Note that we have
$$S_q(X) \leq S_q(Y) \Longleftrightarrow S_q(X|Y) \leq S_q(Y|X) \tag{37}$$

from the symmetry of Tsallis mutual entropy. We also define the Tsallis conditional mutual entropy
$$I_q(X;Y|Z) \equiv S_q(X|Z) - S_q(X|Y,Z) \tag{38}$$

for three random variables X, Y and Z, and $q > 1$.

In addition, $I_q(X;Y|Z)$ is nonnegative, thanks to Eq.(32). For these quantities, we have the following chain rules.

Theorem 4.3   (1) For three random variables X, Y and Z, and $q > 1$, the chain rule holds:

$$I_q(X;Y,Z) = I_q(X;Z) + I_q(X;Y|Z). \tag{39}$$

10

(2) For random variables $X_1, \cdots, X_n$ and Y, the chain rule holds:

$$I_q(X_1, \cdots, X_n; Y) = \sum_{i=1}^{n} I_q(X_i; Y | X_1, \cdots, X_{i-1}). \tag{40}$$

(Proof) (1) follows from the chain rules of Tsallis entropy.

$$\begin{aligned}
I_q(X; Y | Z) &= S_q(X | Z) - S_q(X | Y, Z) \\
&= S_q(X | Z) - S_q(X) + S_q(X) - S_q(X | Y, Z) \\
&= -I_q(X; Z) + I_q(X; Y, Z).
\end{aligned}$$

(2) follows from the application of Eq.(14) and Eq.(13).

$$\begin{aligned}
I_q(X_1, \cdots, X_n; Y) &= S_q(X_1, \cdots, X_n) - S_q(X_1, \cdots, X_n | Y) \\
&= \sum_{i=1}^{n} S_q(X_i | X_{i-1}, \cdots, X_1) - \sum_{i=1}^{n} S_q(X_i | X_{i-1}, \cdots, X_1, Y) \\
&= \sum_{i=1}^{n} I_q(X_i; Y | X_1, \cdots, X_{i-1}).
\end{aligned}$$

∎

We have the following inequality for Tsallis mutual entropies by the strong subadditivity.

Proposition 4.4 For $q > 1$, we have

$$I_q(X; Z) \leq I_q(X, Y; Z).$$

5 Correlation coefficients and distances for Tsallis entropies

We discuss on the entropic distance between X and Y by means of Tsallis conditional entropy $S_q(X|Y)$. If $S_q(X|Y) = 0 = S_q(Y|X)$, then X and Y are called *equivalence* and denoted by $X \equiv Y$. Defining an entropic distance between X and Y as

$$d_q(X, Y) \equiv S_q(X|Y) + S_q(Y|X), \tag{41}$$

we have the following proposition.

Proposition 5.1 For $q > 1$, d_q satisfies

(i) $d_q(X, Y) = 0$ if and only if $X \equiv Y$.

(ii) $d_q(X, Y) = d_q(Y, X)$.

(iii) $d_q(X, Z) \leq d_q(X, Y) + d_q(Y, Z)$.

(Proof) (i) and (ii) are clear from the definition of d_q. In the case of $n = 3$ in Proposition 3.9, we have the triangle inequality:

$$S_q(X_3|X_1) \leq S_q(X_2|X_1) + S_q(X_3|X_2).$$

Putting $X_1 = Z, X_2 = Y$ and $X_3 = X$ in the aboce inequality, we have

$$S_q(X|Z) \leq S_q(Y|Z) + S_q(X|Y). \tag{42}$$

11

In addition, exchanging X and Z in the above inequality, we have

$$S_q(Z|X) \leq S_q(Y|X) + S_q(Z|Y).$$

Summing the both sides of two above inequalities, we have the proposition. ∎

$I_q(X;Y)$ represents a kind of correlation between X and Y. Following [25], we define a parametrically extended correlation coefficient in terms of Tsallis mutual entropy such that

$$\rho_q(X,Y) \equiv \frac{I_q(X;Y)}{S_q(X,Y)} \tag{43}$$

for $S_q(X) > 0$, $S_q(Y) > 0$ and $q > 1$. Then we have the following proposition.

Proposition 5.2 For $q > 1$, $S_q(X) > 0$ and $S_q(Y) > 0$, we have the following properties.

(i) $\rho_q(X,Y) = \rho_q(Y,X)$.

(ii) $0 \leq \rho_q(X,Y) \leq 1$.

(iii) $\rho_q(X,Y) = 0$ if and only if X and Y are independent each other and $q = 1$.

(iv) $\rho_q(X,Y) = 1$ if and only if $X \equiv Y$.

(Proof) (i) is clear from the definition of ρ_q. (ii) follows from $0 \leq I_q(X;Y) \leq S_q(X) \leq S_q(X,Y)$. (iii) follows from Proposition 3.1. (iv) follows from $I_q(X;Y) = S_q(X,Y) \iff S_q(X|Y) = S_q(Y|X) = 0$. ∎

Again, we define an entropic distance between X and Y by

$$\widetilde{d}_q(X,Y) \equiv 1 - \rho_q(X,Y). \tag{44}$$

Then we have the following proposition.

Proposition 5.3 For $q > 1$, $S_q(X) > 0$, $S_q(Y) > 0$ and $S_q(Z) > 0$, we have the following properties.

(i) $\widetilde{d}_q(X,Y) = 0$ if and only if $X \equiv Y$.

(ii) $\widetilde{d}_q(X,Y) = \widetilde{d}_q(Y,X)$.

(iii) $\widetilde{d}_q(X,Z) \leq \widetilde{d}_q(X,Y) + \widetilde{d}_q(Y,Z)$.

(Proof) (i) and (ii) follows from (iv) and (i) of Proposition 5.2, respectively. From chain rules and Eq.(42), we have

$$
\begin{aligned}
\frac{S_q(X|Z)}{S_q(X,Z)} &= \frac{S_q(X|Z)}{S_q(X|Z) + S_q(Z)} \\
&\leq \frac{S_q(X|Y) + S_q(Y|Z)}{S_q(X|Y) + S_q(Y|Z) + S_q(Z)} \\
&= \frac{S_q(X|Y)}{S_q(X|Y) + S_q(Y|Z) + S_q(Z)} + \frac{S_q(Y|Z)}{S_q(X|Y) + S_q(Y|Z) + S_q(Z)} \\
&\leq \frac{S_q(X|Y)}{S_q(X|Y) + S_q(Z)} + \frac{S_q(Y|Z)}{S_q(Y|Z) + S_q(Z)} \\
&= \frac{S_q(X|Y)}{S_q(X,Y)} + \frac{S_q(Y|Z)}{S_q(Y,Z)}. \tag{45}
\end{aligned}
$$

12

Exchanging X and Z, we have

$$\frac{S_q(Z|X)}{S_q(Z,X)} \leq \frac{S_q(Z|Y)}{S_q(Z,Y)} + \frac{S_q(Y|X)}{S_q(Y,X)}. \tag{46}$$

Summing the both sides of two inequalities Eq.(45) and Eq.(46), we have (iii), since Tsallis joint entropy is symmetry.

∎

We also define another correlation coefficient in terms of Tsallis mutual entropy by

$$\hat{\rho}_q(X,Y) \equiv \frac{I_q(X;Y)}{max\{S_q(X), S_q(Y)\}} \tag{47}$$

for $S_q(X) > 0$, $S_q(Y) > 0$ and $q > 1$. See [26]. Then we have the following proposition.

Proposition 5.4 For $q > 1$, $S_q(X) > 0$ and $S_q(Y) > 0$, we have the following properties.

(i) $\hat{\rho}_q(X,Y) = \hat{\rho}_q(Y,X)$.

(ii) $0 \leq \hat{\rho}_q(X,Y) \leq 1$.

(iii) $\hat{\rho}_q(X,Y) = 0$ if and only if X and Y are independent each other and $q = 1$.

(iv) $\hat{\rho}_q(X,Y) = 1$ if and only if $X \equiv Y$.

(Proof) (i)-(iv) follow straightforwardly.

∎

We also define an entropic distance between X and Y by

$$\hat{d}_q(X,Y) \equiv 1 - \hat{\rho}_q(X,Y). \tag{48}$$

Then we have the following proposition.

Proposition 5.5 For $q > 1$, $S_q(X) > 0$, $S_q(Y) > 0$ and $S_q(Z) > 0$, we have the following properties.

(i) $\hat{d}_q(X,Y) = 0$ if and only if $X \equiv Y$.

(ii) $\hat{d}_q(X,Y) = \hat{d}_q(Y,X)$.

(iii) $\hat{d}_q(X,Z) \leq \hat{d}_q(X,Y) + \hat{d}_q(Y,Z)$.

(Proof) (i) and (ii) follows from Proposition 5.4. (iii) is proven by the similar way in [26]. In order to show (iii), we can assume $S_q(Z) \leq S_q(X)$ without loss of generality. Then we prove (iii) in the following three cases.

(a) $S_q(Y) \leq S_q(Z) \leq S_q(X)$: From Eq.(37) and Eq.(42), we have

$$\hat{d}_q(X,Z) = \frac{S_q(X|Z)}{S_q(X)} \leq \frac{S_q(X|Y)}{S_q(X)} + \frac{S_q(Y|Z)}{S_q(X)} \leq \frac{S_q(X|Y)}{S_q(X)} + \frac{S_q(Y|Z)}{S_q(Z)}$$

$$\leq \frac{S_q(X|Y)}{S_q(X)} + \frac{S_q(Z|Y)}{S_q(Z)} = \hat{d}_q(X,Y) + \hat{d}_q(Y,Z).$$

(b) $S_q(Z) \leq S_q(X) \leq S_q(Y)$: It is shown as similar as (a).

13

(c) $S_q(Y) \leq S_q(Z) \leq S_q(X)$: From Eq.(42), we have

$$
\begin{aligned}
\hat{d}_q(X, Z) &= \frac{S_q(X|Z)}{S_q(X)} = 1 - \left(1 - \frac{S_q(X|Z)}{S_q(X)}\right) \leq 1 - \frac{S_q(X)}{S_q(Y)}\left(1 - \frac{S_q(X|Z)}{S_q(X)}\right) \\
&= \frac{S_q(Y) - S_q(X) + S_q(X|Z)}{S_q(Y)} \leq \frac{S_q(Y) - S_q(X) + S_q(X|Y) + S_q(Y|Z)}{S_q(Y)} \\
&= \hat{d}_q(X, Y) + \hat{d}_q(Y, Z).
\end{aligned}
$$

■

6 Applications of Tsallis entropies

6.1 Generalized Fano's inequality

As an application of Tsallis conditional entropy, we give a generalized Fano's inequality.

Theorem 6.1 For $q > 1$ and two random variables taking values in the finite (alphabet) set $\mathcal{X}$, the following inequality holds.

$$
S_q(X|Y) \leq P(X \neq Y)\ln_q(|\mathcal{X}| - 1) + s_q(P(X \neq Y)),
$$

where s_q represents the Tsallis binary entropy:

$$
s_q(p) \equiv -p^q \ln_q p - (1 - p)^q \ln_q(1 - p).
$$

(Proof) Define the random variables Z by

$$
Z = \begin{cases} 0 & \text{if } X = Y \\ 1 & \text{if } X \neq Y. \end{cases}
$$

By chain rules, $S_q(Z|X, Y) = 0$ and $S_q(Z|Y) \leq S_q(Z) = s_q(P(X \neq Y))$, we then have

$$
\begin{aligned}
S_q(X|Y) &= S_q(X|Y) + S_q(Z|X, Y) \\
&= S_q(X, Z|Y) \\
&= S_q(X|Y, Z) + S_q(Z|Y) \\
&\leq S_q(X|Y, Z) + S_q(Z) \\
&= S_q(X|Y = y, Z = 0) + S_q(X|Y = y, Z = 1) + s_q(P(X \neq Y)) \\
&= S_q(X|Y = y, Z = 1) + s_q(P(X \neq Y)) \\
&\leq \ln_q(X \neq Y) + s_q(P(X \neq Y)),
\end{aligned}
$$

since the nonnegativity of Tsallis relative entropy implies $-\sum_{i=1}^{n} p_i^q \ln_q p_i \leq \ln_q n$.

■

6.2 Entropy rate for Tsallis entropy

Moreover, we discuss on the Tsallis entropy rate.

Definition 6.2 The Tsallis entropy rate of a stochastic process $\mathbf{X} \equiv \{X_i\}$ is defined by

$$
S_q(\mathbf{X}) \equiv \lim_{n \to \infty} \frac{S_q(X_1, \cdots, X_n)}{n}
$$

if the limit exists.

$S_q(\mathbf{X})$ can be considered as Tsallis entropy per a symbol in the limit.

Theorem 6.3 For $q > 1$ and a stationary stochastic process, the Tsallis entropy rate exsits and is expressed by

$$S_q(\mathbf{X}) = \lim_{n \to \infty} S_q(X_n | X_1, \cdots, X_{n-1}).$$

(Proof)
By the stationary process and the inequality Eq.(33), we have

$$
\begin{aligned}
S_q(X_n | X_1, \cdots, X_{n-1}) &= S_q(X_{n+1} | X_2, \cdots, X_n) \\
&\geq S_q(X_{n+1} | X_1, X_2, \cdots, X_n).
\end{aligned}
$$

Thus $S_q(X_n | X_1, \cdots, X_{n-1})$ converges to a certain value s, since it is a monotone decreasing sequence with respect to n. That is,

$$\lim_{n \to \infty} S_q(X_n | X_1, \cdots, X_{n-1}) = s.$$

Moreover by the chain rule and Cesáro's theorem (e.g., see pp.64 of [22]), we have

$$S_q(\mathbf{X}) \equiv \lim_{n \to \infty} \frac{S_q(X_1, \cdots, X_n)}{n} = \lim_{n \to \infty} \sum_{i=1}^{n} \frac{S_q(X_i | X_1, \cdots, X_{i-1})}{n} = s.$$

■

6.3 Generalized data processing inequality

Finally we prove the data processing inequality for Tsallis mutual entropy. Let X, Y and Z be random variables. If the conditional probability distribution of the random variable Z depends only on the random variable Y, namely the random variable Z is independent of the random variable X, then it is called that X, Y and Z form a simple Markov chain, and denoted by $X \to Y \to Z$.

Theorem 6.4 If $X \to Y \to Z$, then we have the inequality

$$I_q(X; Z) \leq I_q(X; Y) \tag{49}$$

for $q > 1$, with equality if and only if $X \to Z \to Y$.

(Proof) From Eq.(39), $I_q(X; Y, |Z)$ can be expanded in the following two ways.

$$I_q(X; Y) + I_q(X; Z | Y) = I_q(X; Y, Z) = I_q(X; Z) + I_q(X; Y | Z). \tag{50}$$

Since the nonnegativity $I_q(X; Y | Z)$ and $I_q(X; Z | Y) = 0$ because of the Markovity, we have $I_q(X; Z) \leq I_q(X; Y)$. If $I_q(X; Y | Z) = 0$, namely $X \to Z \to Y$, then we have $I_q(X; Z) = I_q(X; Y)$.

■

We easily find that the following corollary holds.

Corollary 6.5 (1) If $Z = f(Y)$ with a certain mapping f, then $I_q(X; f(Y)) \leq I_q(X; Y)$, with equality if and only if f is a one-to-one mapping.

(2) If $X \to Y \to Z$, then $I_q(X; Y | Z) \leq I_q(X; Y)$.

7 Remarks and discussions

7.1 Chain rule and subadditivity for nonadditive entropies

In [7], the nonadditive entropies including Tsallis entropy and the entropy of type β was axiomatically characterized by the generalized Shannon-Khinchin's axiom in the following manner. The function $\widetilde{S}_q(x_1, \cdots, x_n)$ is assumed to be defined for the n-tuple $(x_1, \cdots, x_n)$ belonging to

$$\Delta_n \equiv \left\{ (p_1, \cdots, p_n) \Big| \sum_{i=1}^n p_i = 1, p_i \geq 0 \ (i = 1, 2, \cdots, n) \right\}$$

and to take a value in $\mathrm{R}^+ \equiv [0, \infty)$. Then it was shown in [7] that four conditions *continuity*, *maximality*, *generalized Shannon additivity* and *expandability* in [7] determine the function $\widetilde{S}_q : \Delta_n \to \mathrm{R}^+$ such that

$$\widetilde{S}_q(x_1, \cdots, x_n) = \frac{\sum_{i=1}^n (x_i^q - x_i)}{\phi(q)}, \tag{51}$$

where $\phi(q)$ is characterized by the following conditions:

(i) $\phi(q)$ is continuous and $\phi(q)(1 - q) > 0$ for $q \neq 1$.

(ii) $\lim_{q \to 1} \phi(q) = 0$ and $\phi(q) \neq 0$ for $q \neq 1$.

(iii) There exists $(a, b) \subset \mathrm{R}^+$ such that $a < 1 < b$ and $\phi(q)$ is differentiable on $(a, 1)$ and $(1, b)$.

(iv) There exists a positive constant number k such that $\lim_{q \to 1} \frac{d\phi(q)}{dq} = -\frac{1}{k}$.

Simple calculations show the nonextensivity

$$\widetilde{S}_q(X \times Y) = \widetilde{S}_q(X) + \widetilde{S}_q(Y) + \phi(q)\widetilde{S}_q(X)\widetilde{S}_q(Y)$$

for the nonadditive entropy $\widetilde{S}_q$. From the view of the generalization Eq.(51), we have the following chain rule.

Lemma 7.1 For the nonadditive entropy

$$\widetilde{S}_q(X) \equiv \frac{\sum_x (p(x)^q - p(x))}{\phi(q)}$$

and any continuous function $\phi(q)$, we have the chain rule

$$\widetilde{S}_q(X, Y) = \widetilde{S}_q(X) + \widetilde{S}_q(Y|X),$$

where the nonadditive joint entropy and the nonadditive conditional entropy are defined by

$$\widetilde{S}_q(X, Y) \equiv \frac{\sum_{x,y} (p(x, y)^q - p(x, y))}{\phi(q)}$$

and

$$\widetilde{S}_q(X|Y) \equiv \sum_y p(y)^q \widetilde{S}_q(X|y) = \sum_y p(y)^q \left\{ \frac{\sum_x (p(x|y)^q - p(x|y))}{\phi(q)} \right\},$$

respectively.

(**Proof**) It follows immediately.

$$\begin{aligned}
\widetilde{S}_q(Y|X) &= \frac{\sum_{y,x} p(x)^q p(y|x)^q - \sum_x p(x)^q}{\phi(q)} \\
&= \frac{\sum_{x,y} (p(x,y)^q - p(x,y))}{\phi(q)} - \frac{\sum_x (p(x)^q - p(x))}{\phi(q)} \\
&= \widetilde{S}_q(X,Y) - \widetilde{S}_q(X).
\end{aligned}$$

∎

From the view of the generalization Eq.(51), Proposition 3.1 also generalized as follows.

Lemma 7.2 For the nonadditive entropy $\widetilde{S}_q(X)$ where $\phi(q)$ is continuous and $\phi(q)(1-q) > 0$ for $q \neq 1$, we have the inequality :

$$\widetilde{S}_q(X|Y) \leq \widetilde{S}_q(X).$$

(**Proof**) Since the function $\widetilde{\mathrm{Ln}}_q(x) \equiv \frac{x^q - x}{\phi(q)}, (q \geq 0, q \neq 1, x \geq 0)$ is nonnegative for $0 \leq x \leq 1, q > 1$ and concave in x for $q > 0, q \neq 1$, the theorem follows as similar as the proof of Proposition 3.1.

∎

Form Lemma 7.1 and Lemma 7.2, we have the subadditivity for nonadditive entropy.

Theorem 7.3 (Subadditivity for nonadditive entropy) For the nonadditive entropy $\widetilde{S}_q(X)$ where $\phi(q)$ is continuous and $\phi(q)(1-q) > 0$ for $q \neq 1$, we have the subadditivity :

$$\widetilde{S}_q(X,Y) \leq \widetilde{S}_q(X) + \widetilde{S}_q(Y).$$

We note that we need the condition that $\phi(q)(1-q) > 0$ for $q \neq 1$ to prove Lemma 7.2, while we do not need any condition to prove Lemma 7.1.

7.2 Inequalities on pseudoadditivity

The pseudoadditivity Eq.(3) for independent random variables X and Y gives rise to the mathematical interest whether we have the complete ordering between the left hand side and the right hand side in Eq.(3) for two general random variables X and Y. Such a kind of problem was taken in the paper [28] for the normalized Tsallis type entropies which are different from the definitions of the Tsallis type entropies in the present paper. However, its inequality appeared in (3.5) of the paper [28] was not true as we found the counter example in [29].

Unfortunately, in the present case, we also find the counter example for the inequalities between $S_q(X,Y)$ and $S_q(X) + S_q(Y) + (1-q)S_q(X)S_q(Y)$. In the same setting of Eq.(27) and Eq.(28), $\delta \equiv S_q(X,Y) - \{S_q(X) + S_q(Y) + (1-q)S_q(X)S_q(Y)\}$ takes both positive and negative values for both cases $0 \leq q < 1$ and $q > 1$. Indeed, $\delta = 0.00846651$ when $q = 1.8, p = 0.1, x = 0.1, y = 0.8$, while $\delta = -0.0118812$ when $q = 1.8, p = 0.1, x = 0.8, y = 0.1$. Also, $\delta = 0.00399069$ when $q = 0.8, p = 0.1, x = 0.8, y = 0.1$, while $\delta = -0.0128179$ when $q = 0.8, p = 0.1, x = 0.1, y = 0.8$.

Therefore there does not exist the complete ordering between $S_q(X,Y)$ and $S_q(X) + S_q(Y) + (1-q)S_q(X)S_q(Y)$ for both cases $0 \leq q < 1$ and $q > 1$.

7.3 A remarkable inequality derived from subadditivity for Tsallis entropies

From Eq.(25), we have the following inequality

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{m}p_{ij}\right)^{r}+\sum_{j=1}^{m}\left(\sum_{i=1}^{n}p_{ij}\right)^{r}\leq\sum_{i=1}^{n}\sum_{j=1}^{m}p_{ij}^{r}+\left(\sum_{i=1}^{n}\sum_{j=1}^{m}p_{ij}\right)^{r} \tag{52}$$

for $r \geq 1$ and p_{ij} satisfying $0 \leq p_{ij} \leq 1$ and $\sum_{i=1}^{n}\sum_{j=1}^{m}p_{ij} = 1$. By putting $p_{ij} = \frac{a_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}}$ in Eq.(52), we have the following inequality as a corollary of Theorem 3.2.

Corollary 7.4 For $r \geq 1$ and $a_{ij} \geq 0$,

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{m}a_{ij}\right)^{r}+\sum_{j=1}^{m}\left(\sum_{i=1}^{n}a_{ij}\right)^{r}\leq\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}^{r}+\left(\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}\right)^{r}. \tag{53}$$

It is remarkable that the following inequality holds [27]

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{m}a_{ij}\right)^{r}\geq\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}^{r} \tag{54}$$

for $r \geq 1$ and $a_{ij} \geq 0$.

7.4 Difference between Tsallis entropy and Shannon entropy

We point out on the difference between Tsallis entropy and Shannon entropy from the viewpoint of mutual entropy. In the case of $q = 1$, the relative entropy between the joint probability $p(x, y)$ and the direct probability $p(x)p(y)$ is equal to the mutual entropy:

$$D_{1}((X,Y)|X\times Y)=S_{1}(X)-S_{1}(X|Y).$$

However, in the general case ($q \neq 1$), there exists the following relation:

$$D_{q}((X,Y)|X\times Y)=S_{q}(X)-S_{q}(X|Y)$$
$$+\sum_{i,j}p(x_{i},y_{j})\left\{p(x_{i})^{q-1}\ln_{q}p(x_{i})+p(y_{j})^{q-1}\ln_{q}p(y_{j})-p(x_{i},y_{j})^{q-1}\ln_{q}p(x_{i})p(y_{j})\right\}, \tag{55}$$

which gives the crucial difference between the special case ($q = 1$) and the general case ($q \neq 1$). The third term of the right hand side in the above equation Eq.(55) vanishes if $q = 1$. The existence of the third term of Eq.(55) means that we have two possibilities of the definition of Tsallis mutual entropy, that is, $I_{q}(X;Y)\equiv S_{q}(X)-S_{q}(X|Y)$ or $I_{q}(X;Y)\equiv D_{q}((X,Y)|X\times Y)$. We have adopted the former definition in the present paper, along with the definition of the capacity in the origin of information theory by Shannon [24].

7.5 Another candidate of Tsallis conditional entropy

It is remarkable that Tsallis entropy $S_{q}(X)$ can be regarded as the expected value of $\ln_{q}\frac{1}{p(x)}$, that is, since $\ln_{q}(x) = -x^{1-q}\ln_{q}(1/x)$, it is expressed by

$$S_{q}(X)=\sum_{x}p(x)\ln_{q}\frac{1}{p(x)}, \quad (q\neq 1), \tag{56}$$

where the convention $0\ln_{q}(\cdot) = 0$ is set. Along with the view of Eq.(56), we may define Tsallis conditional entropy and Tsallis joint entropy in the following.

Definition 7.5 For the conditional probability $p(x|y)$ and the joint probability $p(x,y)$, we define Tsallis conditional entropy and Tsallis joint entropy by

$$\hat{S}_q(X|Y) \equiv \sum_{x,y} p(x,y) \ln_q \frac{1}{p(x|y)}, \quad (q \neq 1), \tag{57}$$

and

$$S_q(X,Y) \equiv \sum_{x,y} p(x,y) \ln_q \frac{1}{p(x,y)}, \quad (q \neq 1). \tag{58}$$

We should note that Tsallis conditional entropy defined in Eq.(57) is not equal to that defined in Eq.(4), while Tsallis joint entropy defined in Eq.(58) is equal to that defined in Eq.(5). If we adopt the above definitions Eq.(57) instead of Eq.(4), we have the following inequality.

Proposition 7.6 For $q > 1$, we have

$$S_q(X,Y) \leq S_q(X) + \hat{S}_q(Y|X).$$

For $0 \leq q < 1$, we have

$$S_q(X,Y) \geq S_q(X) + \hat{S}_q(Y|X).$$

Therefore we do not have the chain rule for $\hat{S}_q(Y|X)$ in general, namely we are not able to construct a parametrically extended information theory in this case.

8   Conclusion

As we have seen, we have found that the chain rules for the Tsallis type entropies hold as similar way of the proofs of the entropy of type β. This result was generalized from a few insights. Also we have proved the strong subadditivity of the Tsallis type entropies for the general two random variables in the case of $q > 1$. Moreover it has been shown that in general the superadditivity of Tsallis entropies does not hold in the case of $0 \leq q < 1$. Thus we could give important results for Tsallis entropies in the case of $q > 1$ from the information theoretical point of view. Therefore we have the possibility that the parametrically extended information theory will be constructed by starting from Tsallis entropy for $q > 1$. In other words, there is less possibility of such a construction in the case of $0 \leq q < 1$ if our stage still in the usual probability theory, because the relation which should be satisfied in the contrast to the usual information theory does not hold in the case of $0 \leq q < 1$. This gives us the rough meaning of the parameter q. That is, we found the crucial different results between $S_q(X)$ for $q > 1$ and $S_q(X)$ for $0 \leq q < 1$ from the information theoretical point of view. For $q > 1$, we showed that the Tsallis type entropies have the similar results to Shannon entropies. In other words, it has been shown that the definition of Tsallis conditional entropy Eq.(4) and Tsallis mutual entropy Eq.(36) have a possibility to be a pioneer to construct a nonadditive information theory.

Our results in the present paper are fundamental so that we are convinced that these results will help the progresses of nonextensive statistical physics and information theory. We are now studying the coding problem in nonadditive information theory and searching for the precise meaning of the parameter q.

Acknowledgement

References

[1] C.Tsallis, Possible generalization of Bolzmann-Gibbs statistics, J.Stat.Phys.,Vol.52,pp.479-487(1988).

[2] C. Tsallis et al., Nonextensive Statistical Mechanics and Its Applications, edited by S. Abe and Y. Okamoto (Springer-Verlag, Heidelberg, 2001); see also the comprehensive list of references at http://tsallis.cat.cbpf.br/biblio.htm.

[3] C. Tsallis, M. Gell-Mann, Y. Sato, "Asymptotically scale-invariant occupancy of phase space makes the entropy S_q extensive", Proc.Nat.Acad.Sci. Vol.102,pp.15377-15382(2005).

[4] C. Tsallis, "Is the entropy S_q extensive or nonextensive?", cond-mat/0409631.

[5] R.J.V. dos Santos, Generalization of Shannon's theorem for Tsallis entropy, J.Math.Phys.,Vol.38,pp.4104-4107(1997).

[6] S.Abe, Axioms and uniqueness theorem for Tsallis entropy, Phys.Lett.A,Vol.271,pp.74-79(2000).

[7] H.Suyari, Generalization of Shannon-Khinchin axioms to nonextensive systems and the uniqueness theorem for the nonextensive entropy, IEEE Trans. on Information Theory, Vol.50, pp.1783-1787 (2004).

[8] S.Furuichi, On uniqueness theorem for Tsallis entropy and Tsallis relative entropy, IEEE Trans. on Information Theory Vol.51, pp.3638-3645 (2005).

[9] L.Borland, A.R.Plastino and C.Tsallis, Information gain within nonextensive thermostatistics, J.Math.Phys,Vol.39,pp.6490-6501(1998), and its Erratum, Vol.40,pp.2196(1999).

[10] S.Abe and A.K.Rajagopal, Nonadditive conditional entropy and its significance for local realism, Physica A, Vol.289, pp.157-164(2001).

[11] T.Yamano, Information theory based on nonadditive information content, Phys.Rev.E,Vol.63,046105(2001).

[12] H.Suyari, Nonextensive entropies derived from form invariance of pseudoadditivity, Phys.Rev.E., Vol. 65, 066118 (2002).

[13] S.Abe, Stability of Tsallis entropy and instabilities of Rényi and normalized Tsallis entropies: A basis for q-exponential distributions, Phys. Rev. E, Vol. 66, 046134(2002).

[14] A.Rényi, On measures of entropy and information, Proc.Forth Berkeley Symp.Math.Stat.Prob., Vol.1,pp.547-561, Univ.California Press,Berkeley, 1961.

[15] A.Rényi, Wahrscheinlichkeitsrechnung. Mit einem Anhang über Informationsteorie., VEB Deutscher Varlan der Wissenschaften, Berlin, 1962.

[16] J.Aczél and Z.Daróczy, On measures of information and their characterizations, Academic Press, 1975.

[17] Z.Daróczy, General information functions, Information and Control, Vol.16,pp.36-51(1970).

[18] J.H.Havrda and F.Charvát, Quantification method of classification processes : Concept of structural a- entropy, Kybernetika (Prague), Vol.3,pp.30-35(1967).

[19] C.Tsallis, Generalized entropy-based criterion for consistent testing, Phys.Rev.E,Vol.58,pp.1442-1445(1998).

[20] M.Shiino, H-theorem with generalized relative entropies and the Tsallis statistics, J.Phys.Soc.Japan,Vol.67,pp.3658-3660(1998).

[21] A.K.Rajagopal and S.Abe,Implications of form invariance to the structure of nonextensive entropies, Phys. Rev. Lett. Vol.83, pp.1711-1714(1999).

[22] T.M.Cover and J.A.Thomas, Elements of Information Theory, John Wiley and Sons, 1991.

[23] S.Furuichi, K.Yanagi and K.Kuriyama, Fundamental properties of Tsallis relative entropy, J.Math.Phys,Vol.45,pp.4868-4877(2004).

[24] C.E.Shannon, A mathematical theory of communication, The Bell System Technical Journal, Vol.27,pp.379-423,pp.623-656,(1948).

[25] C.Rajski, A metric space of discrete probability distributions, Information and Control, Vol.4,pp.371-377(1961).

[26] Y.Horibe, Information entropy theory, Morikita publishing company (in Japanese), 1989.

[27] G.H.Hardy, J.E.Littlewood and G.Pólya, Inequality, Cambridge University Press, 1952.

[28] K.S.Fa, Note on generalization of Shannon theorem and inequality, J.Phys.A; Math. and Gen.,Vol.31,pp.8159-8164(1998).

[29] H.Suyari and S.Furuichi, Comment on "Note on generalization of Shannon theorem and inequality", J.Phys.A: Math. and Gen., Vol.36,pp.5157-5158(2003).